DOCUMENT RESUME

ED 113 972 FL 007 283

TITLE	Compte Rendu de Recherche sur le Programme Lexical. Langue Anglaise 6e et 5e (Research Report on the Lexical Program. English Language 5th and 6th).
INSTITUTION	Institut National de Recherche et de Documentation Pedagogiques, Paris (France).
PUB DATE	73
NOTE	50p.; In French
AVAILABLE FROM	Institut National de Recherche et de Documentation Pedagogiques, 29 rue d'Ulm, 75 Paris 5, France (6,50F)
EDRS PRICE	MF-$0.76 HC-$1.95 Plus Postage
DESCRIPTORS	*Basic Vocabulary; *English (Second Language); *French; Language Fluency; *Language Instruction; Language Research; Language Skills; Native Speakers; Research Methodology; Second Language Learning; *Vocabulary Development; Vocabulary Skills

ABSTRACT

This research report is concerned with establishing a minimum vocabulary of English instruction for French students at the secondary level. The approach to the investigation is based on choosing vocabulary that corresponds to the current communication needs of the students, has functional content, and takes the spoken language into account. The need for limiting the volume of the vocabulary is stressed. Two basic criteria are defined for word selection. The criterion of function and distribution considers such things as word frequency, number and kinds of contexts a word appears in, grammatical function, cultural content, and the vocabulary needed for the language learning situation. The criterion of facility in language learning considers French cognates a word may have, and a word's regularity, simplicity and semantic precision. Four methods of choosing and classifying vocabulary items are described and compared. The first appendix lists the scientific works consulted, while the second appendix provides a sample index card used in the study. The third appendix lists the English function words to be included in the vocabulary. (CLK)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED113972

FL007283

FV

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
INRDP
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

RECHERCHES PÉDAGOGIQUES

INSTITUT NATIONAL DE RECHERCHE ET DE DOCUMENTATION PÉDAGOGIQUES

9
PÉRIODIQUES
12
COLLECTIONS
AU SERVICE DE L'ENSEIGNEMENT

catalogue gratuit
sur
demande

INSTITUT NATIONAL
DE RECHERCHE
ET DE DOCUMENTATION
PÉDAGOGIQUES

SERVICE D'ÉDITION ET DE VENTE
DES PUBLICATIONS
DE L'ÉDUCATION NATIONALE

29 RUE D'ULM - 75-PARIS-5

SEV 2

INSTITUT NATIONAL DE RECHERCHE ET DE DOCUMENTATION PÉDAGOGIQUES

SERVICE DES ETUDES ET RECHERCHES PEDAGOGIQUES

COMPTE RENDU DE RECHERCHE sur le PROGRAMME LEXICAL

LANGUE ANGLAISE
6e et 5e

Brochure n° 59 RP

1973

Dans la même collection

L'enseignement programmé en latin et grammaire (1968) 32 RP

Les mathématiques en 5e - Expérimentation et nouveaux programmes (1970) 42 RP

Recherches dans l'enseignement du français au second degré - I (1971) 43 RP

Recherches dans l'enseignement du français au second degré - Grammaire, expression, méthodes - II (1971) 49 RP

Linguistique et enseignement du français - Recherches au niveau du premier cycle - 1969-1970 1972-71 (1972) 52 RP

Linguistique fonctionnelle et enseignement du français 57 RP (à paraître).

Document de recherche
à l'usage des établissements
charges d'expérimentation
Tous droits réservés

SOMMAIRE

Pages

Commission du lexique anglais - Compte rendu de recherche

I - Hypothèses de travail 7

II - Les sources 8

III - Constitution du fichier 8

IV - Méthodes de travail 9

V - Présentation des listes 23

Annexe 1 : — Liste analytique des ouvrages dépouillés 33

Annexe 2 : — Modèle de fiche 41

— Contenu de la fiche 42

Annexe 3 : — Les mots grammaticaux 43

Bibliographie 46

The vocabulary in a well-graded language-course will be arranged in such a manner that the more useful words will be learnt before the less useful. (Let us remember that there are two sorts of « useful words » ; those which are useful in themselves *on account of their intrinsic meaning, and words which are* useful as sentence-formers.)

. . . . . .

If a given word has several meanings, let us first associate the word with its more usual or useful meanings.

Harold E. PALMER

« The Principles of Language-Study »

© 1922 - Reprinted 1964, OUP, London

COMMISSION DU LEXIQUE ANGLAIS
COMPTE RENDU DE RECHERCHE

Le but de la commission a été, dans un premier temps, d'établir un minimum lexical pour l'enseignement de l'anglais au premier cycle à commencer par les classes de 6e et 5e.

Les professeurs d'anglais savent combien il est difficile de faire acquérir aux élèves, tout au long des sept années de l'enseignement secondaire, une connaissance de la langue anglaise aussi complète et efficace que possible.

En particulier il est toujours malaisé de déterminer le volume et la nature des notions qu'il est opportun de faire assimiler aux élèves, année après année.

Un effort avait déjà été fait dans ce sens : le « Programme grammatical pour l'enseignement de l'anglais », qui prévoit pour chaque classe le contenu minimum à enseigner. Les auteurs de ce programme s'étaient donné pour tâche d'organiser et de découper un système morpho-syntaxique.

Si la grammaire d'une langue, tout comme son système phonologique, peut être considérée comme un système fermé, le lexique, au contraire, constitue un ensemble ouvert et beaucoup plus mouvant que la grammaire. On mesure ainsi les difficultés rencontrées dans l'établissement du programme lexical pour l'enseignement de l'anglais.

I - HYPOTHESES DE TRAVAIL

Il a semblé utile d'adopter la démarche suivante :

1) Choisir un contenu fonctionnel, plutôt que descriptif, qui tienne le plus grand compte des exigences de la langue parlée, conformément à l'esprit des méthodes modernes d'enseignement des langues vivantes (1).

2) Choisir un niveau de langue qui corresponde aux besoins de la communication la plus courante et de l'expression personnelle orale et écrite.

---

(1) Voir instructions du 28 août 1969, B.O. 4-9-69.

3) Harmoniser le programme lexical avec le programme grammatical.

4) Associer aux critères linguistiques des critères psycho-pédagogiques, pour tenir compte des possibilités langagières et des intérêts propres aux élèves suivant leur âge.

5) Limiter le nombre des unités lexicales obligatoires et celui de leurs acceptions dans un souci d'efficacité et de réalisme (effectifs et horaires actuels).

II - LES SOURCES

La commission a travaillé sur un ensemble d'unités lexicales extraites :

a) d'ouvrages scientifiques américains, anglais, australiens, français et irlandais qui s'appuient sur des corpus de langue écrite ou de langue parlée (langue des adultes ou langue des enfants) et qui utilisent différents critères de choix : fréquence, répartition, disponibilité, rentabilité structurale, motivation (1) ;

b) du contenu lexical de six manuels français pour les classes de 6e et de 5e, de deux méthodes audio-visuelles et de deux manuels britanniques à l'usage des débutants (2).

Les unités lexicales ainsi recueillies ont été reportées sur fiches.

III - CONSTITUTION DU FICHIER

Afin de pouvoir exploiter ensuite au mieux les sources énumérées ci-dessus, le premier travail de la commission a été d'établir un fichier en prenant comme base de départ la « General Service List » de Michael West.

Il n'a pas échappé à la commission que la conception de cet ouvrage différait sensiblement de ses propres objectifs et de ses principes de travail. On sait en effet que la « G.S.L. » a été établie à partir d'un corpus de langue écrite, que la fréquence a été retenue comme unique critère de choix et que les dérivés sont présentés dans la même rubrique sans être comptabilisés séparément.

On a cependant considéré que, à défaut de données plus adéquates, la « G.S.L. » était l'ouvrage de référence le plus commode pour servir de base

---

(1) Voir annexe I la liste analytique de ces ouvrages.

(2) Voir annexe II le modèle de la fiche utilisée comportant la liste de ces manuels.

à l'établissement du fichier de la commission. Le corpus étudié est en effet suffisamment étendu pour couvrir les besoins d'un programme lexical à l'usage de l'ensemble des classes du 1er cycle et l'indication de fréquence est complétée par des pourcentages sémantiques (calculs de Thorndike-Lorge).

De plus, si la « G.S.L. » a servi de base à l'établissement du fichier, elle n'a pas été l'unique instrument de travail de la commission qui a étudié soigneusement les autres ouvrages scientifiques sans négliger pour autant les manuels indiqués au paragraphe II.

Les fiches ainsi établies ont été examinées séparément et réparties successivement en trois groupes :

1) Les mots (1) qui paraissaient ne devoir être retenus ni pour la liste de 6e ni pour celle de 5e.

2) Les mots qui devaient pouvoir servir de base pour l'établissement de la liste de 6e.

3) Les mots qui, sans être retenus dans un premier temps pour le corpus de 6e, étaient susceptibles de l'être pour celui de 5e.

IV - METHODES DE TRAVAIL

A - CHOIX DES CRITERES

Pour juger des mots, des acceptions ou des emplois à retenir, la commission s'est appuyée sur un système de critères précis qu'elle a tenté de quantifier et d'appliquer de façon mathématique.

1. Liste des critères

a) *Critères d'utilité*

— fréquence

— répartition

— disponibilité (2)

---

(1) C'est par commodité que nous désignons par « mot » les unités lexicales reconnues

(2) Critère défini par R. Michéa dans l'Elaboration du Français Fondamental, cf. annexe 1, p. 37, paragraphe 6.

— Valence (1) composée des sous-critères de :

définition
inclusion
extension
combinaison

— fonctionnement grammatical

— contenu culturel

— mise en œuvre pédagogique.

b) *Critères de facilité d'apprentissage*

— similitude avec la langue maternelle

— régularité

— simplicité

— précision sémantique

— pouvoir motivant.

2. Définition des critères et sous-critères.

Fréquence. — C'est la fréquence relative d'apparition d'un mot dans un corpus donné.

Ex. : « take » plus fréquent (W. : 7 008 e) que « happen » (W. : 790 e).

Répartition. — Elle indique, pour un corpus donné, le nombre de contextes différents dans lesquels le mot apparaît.

Ex. : « go » plus largement réparti qu'un verbe comme « swim » qui n'apparaît que dans un nombre limité de contextes).

Disponibilité. — Les mots « disponibles » sont ceux qui sont indispensables à l'expression dans certaines situations courantes sans avoir nécessairement une fréquence élevée.

Ex. : « fork », mot disponible qui n'a pourtant qu'une fréquence basse (W. : 70 e).

Valence. — C'est un critère d'utilité qui mesure le pouvoir qu'ont les mots de remplacer d'autres mots de 4 façons :

① **Définition.** — Un mot qui, associé à un autre, permet d'en limiter le champ sémantique est considéré comme ayant un pouvoir définissant.

Ex. : « young » associé à « horse », « ox », « sheep », « dog », « cat » permet de définir les mots « colt », « calf », « lamb », « puppy », « kitten », etc.

---

(1) Pour le critère de valence et de ses composants, la commission a bénéficié des travaux effectués depuis plusieurs années par Jean-Guy Savard et Jack Richards sous la direction du Professeur W.-F. Mackey de l'Université de Laval.

② **Inclusion.** — Le pouvoir d'inclusion d'un mot est lié à l'étendue de son champ sémantique.

Ex. : « seat » a un champ sémantique plus étendu, donc un pouvoir d'inclusion plus grand que « chair », « stool » ou « arm-chair ».

③ **Extension.** — Le pouvoir d'extension d'un mot équivaut à son degré de polysémie.

Ex. : le mot « get » qui peut engendrer « get in », « get on », etc., le verbe « make » qui entre dans des expressions comme « make a noise », « make a mistake », etc.

④ **Combinaison.** — Faculté pour un mot de se lier à un autre pour former un mot composé.

Ex. : « foot » (football, footpath, etc.) a un pouvoir de combinaison plus élevé que « leg ».

Fonctionnement grammatical et rentabilité structurale. — C'est le critère qui s'attache de toute évidence aux mots-outils mais aussi à certaines unités lexicales qui se prêtent mieux que d'autres au maniement des structures.

Ex. : 1) L'adjectif long « expensive », l'adjectif court « cheap » ; le substantif « cake » qui appartient à la fois aux dénombrables et aux indénombrables.

2) Verbes de mouvement « come », « go », etc. permettant l'utilisation de certaines prépositions et postpositions.

Contenu culturel. — Il est déterminé par la plus ou moins grande importance des faits de civilisation évoqués par certains mots.

Ex. : « Christmas » a un contenu culturel plus élevé que « Whitsun ».

Mise en œuvre pédagogique. — C'est le vocabulaire exigé par la situation scolaire dans laquelle se déroule l'apprentissage de la langue.

Ex : « chalk », « repeat », « next » (next please).

Similitude avec la langue maternelle. — La ressemblance dans les domaines de la prononciation, de l'orthographe et du contenu sémantique facilite l'apprentissage.

Ex. : « soup » a un très grand degré de similitude avec le français « soupe » (prononciation - orthographe), même contenu sémantique, alors que « large » n'a guère de similaire que l'orthographe.

Régularité. — C'est la plus ou moins grande conformité à un système donné, à l'intérieur de la langue.

Ex. : « sick » opposé à « ill » dans le fonctionnement de l'adjectif.

Simplicité. — Elle dépend de la longueur du mot et des difficultés de prononciation ou d'orthographe.

Ex. : « funny » est plus simple que « exhilarating »,
« second » plus accessible que « twelfth ».

Précision sémantique. — Elle est plus ou moins grande suivant que le référent extra-linguistique est plus ou moins particularisé.

Ex. : « apple » a une précision sémantique plus grande que « fruit », dont le champ sémantique est plus étendu.

Il en est de même pour « beautiful » par rapport à « nice ».

Pouvoir motivant. — Il est lié au rapport entre le mot et les goûts et les intérêts des enfants d'un groupe d'âge déterminé.

Ex. : « car » représente pour les enfants à la fois un élément de jeu et de civilisation.

B - APPLICATION DES CRITERES

Pour tenir compte de l'importance relative des critères, la commission a fait divers essais de classification.

1. Première méthode de travail

Après avoir dressé une liste aussi exhaustive que possible des critères de sélection du vocabulaire, la commission s'est efforcée de leur affecter des coefficients en fonction de leur importance relative.

Pour obtenir commodément des résultats sous forme de pourcentage, les quinze critères retenus ont été affectés chacun d'un coefficient (1, 2 ou 3), la somme de ces coefficients étant de 25. Chaque mot examiné recevait ensuite, pour chacun des critères, une note de 0 à 4 (note maximale éventuelle : 4 × 25 soit 100).

Liste des critères classés dans l'ordre décroissant des coefficients

N° d'ordre	Critère	Coefficient
1	Fréquence	3
2	Fonctionnement grammatical	3
3	Pouvoir motivant	3

4	Répartition	2
5	Disponibilité	2
6	Extension	2
7	Mise en œuvre pédagogique	2
8	Contenu culturel	1
9	Régularité	1
10	Précision sémantique	1
11	Définition	1
12	Simplicité	1
13	Similitude avec la langue maternelle	1
14	Combinaison	1
15	Inclusion	1
	Total	25

Afin de tester la valeur de ce système, la commission a choisi un échantillonnage de mots dans le fichier initial parmi ceux qui, a priori, devaient obtenir un pourcentage significatif.

Par exemple :

Television - a/an - animal - apple - apply.

— Le mot **television** se détache par son pouvoir motivant.

— L'article **a/an** est absolument indispensable dans tout corpus.

— Les substantifs **animal** et **apple** sont des termes courants : le premier est classificateur, le second est spécifique.

— Le verbe **apply** appartient au vocabulaire de l'adulte.

Cette procédure s'est révélée à la fois trop lourde et trop peu discriminative (1) :

1) Le poids de plus de la moitié des critères était négligeable et n'avait qu'une faible influence sur la place du mot dans le classement final.

2) Ce système d'évaluation quantitative n'était pas pleinement satisfaisant. En effet, les coefficients ne jouaient pas de la même manière selon qu'on les appliquait aux mots grammaticaux ou aux mots à sémantisme plein. Les critères à gros coefficient (fréquence, fonctionnement grammatical, disponibilité, répartition et mise en œuvre pédagogique) étaient trop favorables aux mots grammaticaux ; seul le critère de motivation leur était défavorable.

(1) Voir tableau I de la page suivante.

TABLEAU I

	Critères	Coefficient C	television		a - an		animal		apple		apply	
			note	N × C ①	note	N × C	note	N × C	note	N × C	note	N × C
1	Fréquence	3	1	3	4	12	2	6	1	3	1	3
2	Fonctionnement grammatical	3	0	0	4	12	1	3	1	3	1	3
3	Pouvoir motivant .	3	4	12	0	0	2	6	4	12	0	0
4	Répartition	2	0	0	4	8	1	2	1	2	1	2
5	Disponibilité	2	4	8	4	8	4	8	3	6	2	4
6	Extension	2	0	0	4	8	0	0	0	0	0	0
7	Mise en œuvre pédagogique	2	0	0	4	8	0	0	0	0	0	0
8	Contenu culturel ..	1	3	3	0	0	0	0	2	2	0	0
9	Régularité	1	2	2	1	1	4	4	4	4	3	3
10	Précision sémantique	1	4	4	1	1	1	1	4	4	4	4
11	Définition	1	0	0	2	2	4	4	0	0	0	0
12	Simplicité	1	2	2	4	4	3	3	2	2	2	2
13	Similitude	1	3	3	1	1	3	3	0	0	0	0
14	Combinaison	1	0	0	2	2	0	0	2	2	1	1
15	Inclusion	1	0	0	0	0	4	4	0	0	0	0
	Somme des produits (N × C) (2)			37/100		67/100		44/100		40/100		22/100

① N représente la note de 0 à 4 et C le coefficient du critère de 1 à 3

② La somme des produits (N × C) divisée par 15 (nombre de critères) fournit le poids moyen du mot dans l'ensemble des 5 mots comparés. Si les mots avaient tous le même poids, ce poids serait de 20 % ($\frac{100\ \%}{5}$).

20 représente la moyenne théorique.

Les cinq mots se classent alors dans l'ordre du poids relatif en % par rapport à la moyenne théorique 20.

a/an	+ 47 %	(au lieu de 67 %)
animal	+ 24 %	(au lieu de 44 %)
apple	+ 20 %	(au lieu de 40 %)
television	+ 17 %	(au lieu de 37 %)
apply	+ 2 %	(au lieu de 22 %).

2. Deuxième méthode de travail

Afin de disposer d'un outil plus maniable et mieux adapté à son travail, la commission a décidé de limiter à dix le nombre des critères et de supprimer les coefficients.

Nouvelle liste des critères :

1) Fréquence
2) Extension
3) Définition
4) Combinaison
5) Inclusion
6) Fonctionnement grammatical
7) Répartition
8) Disponibilité
9) Pouvoir motivant
10) Mise en œuvre pédagogique.

Pour chaque critère, le mot considéré a reçu la note 1, 2 ou 3.

Afin de tester la valeur de ce deuxième système de notation, la commission a choisi un échantillonnage plus large composé d'unités lexicales plus variées (1).

(1) Voir tableau II de la page suivante.

TABLEAU II

N° d'ordre	Critère	after-noon	a - an	apple	ask	ill	windy	birth-day	bed	big	go	Total obtenu pour chaque critère
1	Fréquence	3	3	1	3	2	1	2	3	3	3	24/30
2	Extension	1	2	1	1	1	1	1	1	1	3	13/30
3	Définition	1	1	1	2	1	1	1	1	3	3	15/30
4	Combinaison	1	1	3	1	1	1	3	1	1	3	16/30
5	Inclusion	1	1	1	2	2	1	1	1	2	3	15/30
6	Fonctionnement grammatical	2	3	1	2	2	1	1	2	3	3	20/30
7	Répartition	3	3	1	3	1	1	1	3	3	3	22/30
8	Disponibilité	3	3	3	3	3	1	2	3	3	3	27/30
9	Pouvoir motivant .	1	1	2	1	2	1	3	3	3	1	18/30
10	Mise en œuvre pédagogique	1	1	1	3	1	1	1	1	1	3	14/30
Total obtenu pour chaque mot		17/30	19/30	15/30	21/30	16/30	10/30	16/30	19/30	23/30	28/30	
Pourcentage		56,7 %	63,3 %	50 %	70 %	53,3 %	33,3 %	53,3 %	63,3 %	76,6 %	93,3 %	

Cette méthode pouvait paraître satisfaisante dans son ensemble. Elle conduisait cependant à un résultat aberrant : le faible total atteint par l'article indéfini (1).

Les résultats obtenus n'offraient pas la garantie d'objectivité souhaitée. En effet, il y avait une part trop grande de subjectivité dans l'évaluation des critères. L'application d'un même système non pondéré à des mots de catégories différentes (mots grammaticaux et mots thématiques) ne se justifiait pas.

3. Troisième méthode de travail

Pour essayer de réduire la part de subjectivité et pour réintroduire une plus grande discrimination entre les catégories de mots, la commission a repris l'ensemble des unités lexicales étudiées selon la première et la deuxième méthodes de travail et les quinze critères retenus à l'origine et a attribué, d'une part, à chaque mot et pour chaque critère, une note de 1 à 5, et d'autre part, un coefficient de 1 à 3 à chaque critère. Le maximum possible de points pour un mot était de 125 (somme des produits des notes maximales par les coefficients). Les résultats obtenus apparaissent dans le tableau suivant (p. 18-19).

---

(1) Ordre relatif des pourcentages obtenus pour chaque mot :

go	15,22 %
big	12,52 %
ask	11,45 %
a/an	10,32 %
bed	10,32 %
afternoon	9,23 %
birthday	8,69 %
ill	8,69 %
apple	8,15 %
windy	5,41 %.

TABLEAU III

Critères	Coefficient	go		a - an		big		ask		bed	
		note	N × C	note	N × C	note	N × C	note	N × C	note	N × C
Fréquence	3	5	15	5	15	5	15	5	15	4	12
Fonctionnement grammatical	3	4	12	5	15	3	9	3	9	2	6
Pouvoir motivant	3	3	9	1	3	4	12	1	3	4	12
Répartition	2	5	10	5	10	5	10	1	2	5	10
Disponibilité	2	5	10	5	10	5	10	5	10	5	10
Extension	2	5	10	5	10	1	2	5	10	1	2
Mise en œuvre pédagogique	2	4	8	5	10	1	2	5	10	1	2
Contenu culturel	1	1	1	1	1	1	1	1	1	1	1
Régularité	1	4	4	2	2	4	4	5	5	5	5
Précision sémantique	1	2	2	2	2	2	2	2	2	3	3
Définition	1	4	4	3	3	5	5	4	4	4	4
Simplicité	1	4	4	5	5	5	5	3	3	5	5
Similitude	1	1	1	2	2	1	1	1	1	1	1
Combinaison	1	2	2	3	3	2	2	1	1	4	4
Inclusion	1	5	5	1	1	3	3	5	5	2	2
Total des coefficients	25		97		92		83		81		79
N × C = 125											
Valeur globale relative			77,6 %		73,6 %		66,4 %		64,8 %		63,2 %

birthday		afternoon		apple		animal		television		ill		apply		windy	
note	N × C	note	N × C	note	N × C	note	N × C	note	N × C	note	N × C	note	N × C	note	N × C
1	3	5	15	2	6	3	9	2	6	2	6	2	6	1	3
3	9	2	6	2	6	2	6	1	3	2	6	2	6	2	6
5	15	1	3	5	15	3	9	5	15	2	6	1	3	1	3
5	10	4	8	2	4	2	4	1	2	1	2	2	4	1	2
2	4	5	10	4	8	2	4	5	10	5	10	3	6	1	2
2	4	1	2	1	2	1	2	1	2	1	2	1	2	1	2
2	4	2	4	1	2	1	2	1	2	2	4	1	2	1	2
2	2	1	1	3	3	1	1	4	4	1	1	1	1	1	1
5	5	5	5	5	5	5	5	3	3	1	1	4	4	1	1
5	5	4	4	5	5	2	2	5	5	1	1	5	5	1	1
1	1	1	1	1	1	5	5	1	1	4	4	1	1	3	3
2	2	1	1	3	3	4	4	3	3	3	3	3	3	4	4
1	1	2	2	1	1	4	4	4	4	1	1	1	1	2	2
1	1	2	2	3	3	1	1	1	1	1	1	2	2	1	1
1	1	1	1	1	1	5	5	1	1	4	4	1	1	3	3
	67		65		65		63		62		52		47		36
	53,6 %		52 %		52 %		50,4 %		49,68 %		41,6 %		37,6 %		28,8 %

La troisième méthode semblait donner de meilleurs résultats. Néanmoins la place de l'article indéfini (en deuxième position) ne paraissait pas non plus refléter suffisamment l'importance qu'il a dans la langue. Dans cette méthode, les mots grammaticaux continuent à être défavorisés par le traitement statistique utilisé, malgré l'application d'un coefficient élevé.

La commission, après avoir procédé à ces trois essais de quantification, en est arrivée à la conclusion que, dans l'état actuel de la recherche linguistique, il lui était impossible, faute de données statistiques et théoriques suffisantes comme de moyens techniques adéquats, d'obtenir des résultats quantifiables d'une valeur statistique incontestable. Elle a donc décidé de se tourner vers une quatrième méthode de travail.

Voir page 21 le tableau comparatif des trois méthodes.

4. Quatrième méthode de travail

Elle consistait, pour chaque mot du corpus, à se servir des 15 critères retenus sans rechercher systématiquement à les appliquer tous et sans vouloir aboutir à des résultats chiffrés.

a - an

— fréquence maximale d'apparition (West : 89 572) (1)

— mot grammatical indispensable

à retenir pour la 6e

afternoon

— rentabilité structurale pour le maniement des temps

— fréquence élevée dans les situations de la vie quotidienne

à retenir pour la 6e

animal

— pouvoir définissant élevé mais faible pouvoir motivant par rapport aux termes plus spécifiques (cat, dog...)

— fréquence West : 849

à ne pas retenir pour la 6e

à retenir pour la 5e

apple

— disponibilité élevée

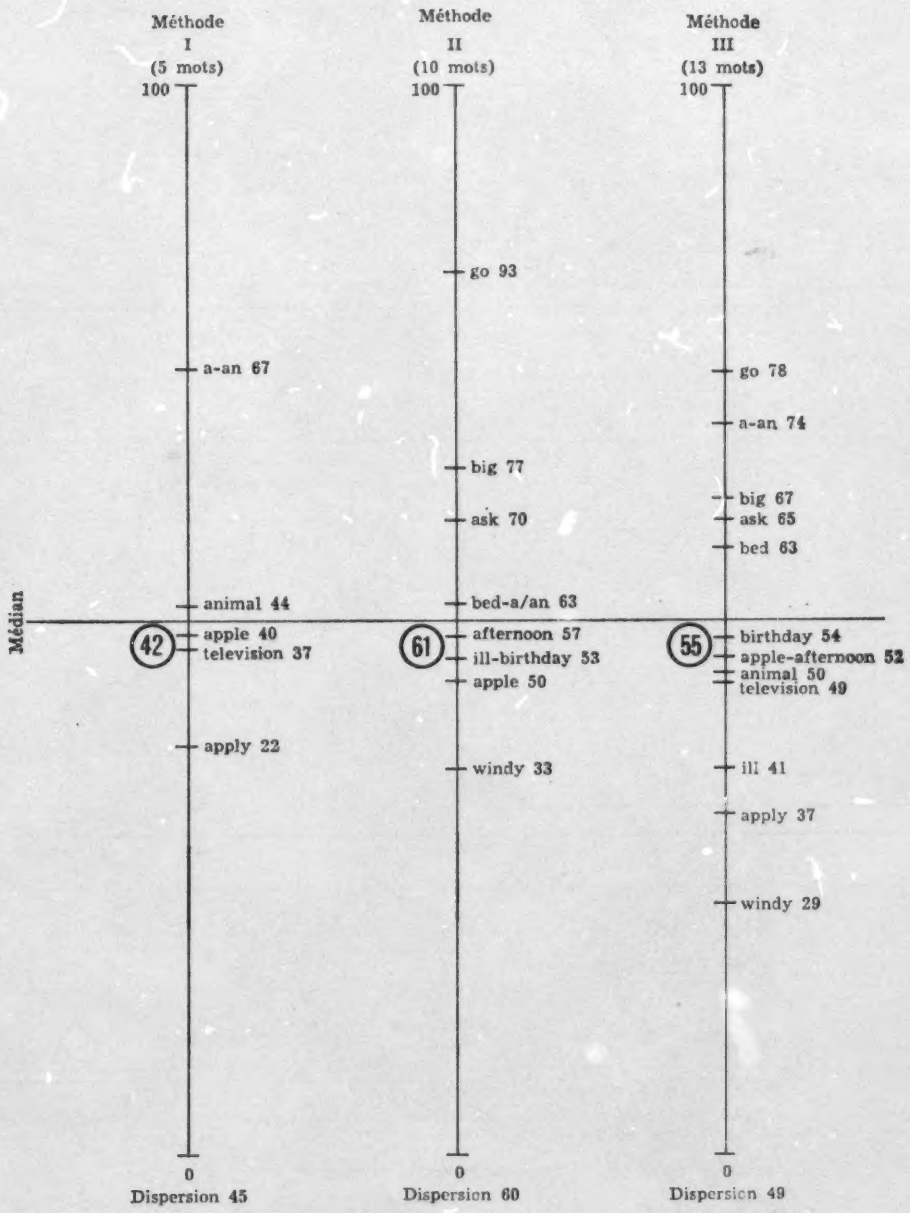
— combinaison facile, grande précision sémantique

— fréquence West : 12

à retenir pour la 6e

(1) La fréquence chez West va de 106 064 à 2.

TABLEAU COMPARATIF DES TROIS MÉTHODES



Place relative, sur une même échelle arithmétique, *autour du médian*, dans les trois méthodes **des écarts entre les mots.**

apply

— peu fréquent (West : 608)

— trop grande difficulté d'apprentissage (interférence sémantique et syntaxique)

— peu motivant pour l'enfant

à ne retenir ni pour la 6e
ni pour la 5e

ask

— utile pour la mise en œuvre pédagogique

— fonctionnement grammatical intéressant :
Ask him a question
to come

— intérêt phonétique (utile entraînement à la prononciation de groupes consonantiques) (asks - asked)

— fréquence élevée (West : 4 062)

à retenir pour la 6e dans ces 2 emplois

bed

— disponibilité élevée

— pouvoir motivant

— combinaison facile

— fréquence West : 904

à retenir pour la 6e

big

— fort pouvoir définissant

— fréquence élevée (West : 1 182)

— fonctionnement grammatical (comparatif - superlatif)

— pouvoir motivant (valeur affective et possibilité de classification)

— degré d'inclusion élevé

à retenir pour la 6e

birthday

— pouvoir motivant pour l'enfant

— combinaison facile

— contenu culturel (civilisation anglo-saxonne)

— fonctionnement grammatical

à retenir pour la 6e

go

— fréquence et répartition très élevées (West : 8 760)

— fonctionnement grammatical

— degrés d'inclusion et d'extension élevés

— mise en œuvre pédagogique

à retenir pour la 6e

ill

— pouvoir motivant élevé

— mais trop grande difficulté d'apprentissage (interférence sémantique et syntaxique)

— fréquence West : 937 e

à ne pas retenir pour la 6e
à retenir pour la 5e

television

— pouvoir motivant très élevé

— facilité d'apprentissage

— contenu culturel (civilisation contemporaine)

à retenir pour la 6e

windy

— fréquence peu élevée : West 14 e

— pouvoir motivant faible

à ne pas retenir ni pour la 6e

ni pour la 5e

Cette dernière méthode de travail a finalement été retenue.

V - PRESENTATION DES LISTES

La commission a jugé utile de distinguer 5 catégories d'unités lexicales.

T : mots thématiques ou à sémantisme plein (noms, verbes et adjectifs, et certains adverbes).

G : mots grammaticaux.

S : séries (nombres, jours de la semaine, mois, etc.).

F : expressions toutes faites (« formulaic utterances »).

P : Pédagogie (mots et expressions utiles dans la conduite de la classe mais qui ne font pas, immédiatement, l'objet d'un enseignement systématique).

A partir du corpus choisi pour la 6e et la 5e, la commission a finalement retenu 1 293 mots et expressions :

	6e	5e	Total
1) mots thématiques ou à sémantisme plein (noms, verbes, adjectifs et certains adverbes) — T	370	489	859
2) mots grammaticaux — G	141	100	241
3) séries (nombres, jours de la semaine, mois, etc.) — S	81	14	95
4) expressions toutes faites (« formulaic utterances ») — F	26	6	32
5) pédagogie (mots et expressions utiles dans la conduite de la classe mais qui ne font pas, immédiatement, l'objet d'un enseignement systématique). — P	51	15	66
	669	624	1 293

Le total des unités retenues représente avec les horaires actuels une moyenne de 5 ou 6 unités lexicales par heure de cours, ce qui correspond aux normes généralement préconisées par les psychologues (1) pour la plupart des élèves.

A - CLASSEMENT DES UNITES

1. Les mots thématiques ou à sémantisme plein (« T »)

a) CONTENU DES LISTES

Dans la perspective des débuts de l'apprentissage de l'anglais, il est évident qu'il fallait considérer en priorité le vocabulaire de la langue parlée, en tenant compte des besoins morpho-syntaxiques à ce niveau. La commission s'est refusée délibérément à choisir les mots en fonction des centres d'intérêt. Un tel classement présente en effet deux dangers :

(1) Voir bibliographie, p. 46.

— par souci d'être exhaustif, on est amené à surcharger inutilement la liste des unités lexicales ;

— on risque d'autre part de perdre de vue le caractère fonctionnel des unités à enseigner.

En revanche, au cours de ses travaux, elle a tenu à s'assurer que les mots retenus correspondaient aux thèmes fondamentaux de la vie de l'enfant et de ses rapports avec le monde des adultes. Il est clair qu'une bonne part du vocabulaire à enseigner aux élèves de 6e et 5e appartient à la fois au monde linguistique des enfants et à celui des adultes.

La commission a surtout pensé à la langue de la communication et de l'expression personnelle dans les situations de la vie courante. Les thèmes fondamentaux qui se sont ainsi dégagés au fur et à mesure appartiennent à trois grandes catégories qui correspondent aux milieux dans lesquels se déroulent les activités de l'enfant : la maison, l'école, l'extérieur.

C'est à ces thèmes fondamentaux et à ces situations que la commission s'est constamment référée en opérant ses choix à l'aide des critères exposés plus haut.

b) LIMITES D'ACCEPTION ET D'EMPLOI. POUVOIR OPERATOIRE DES MOTS.

Un même mot peut très bien se référer au même concept général tout en ayant des **acceptions** ou des **emplois** différents. La commission a jugé bon de ne retenir que certains de ces emplois et acceptions et d'en réserver d'autres pour une étape ultérieure.

Exemple :

1) trois acceptions ont été distinguées pour le substantif « back » : elles sont illustrées par les trois exemples suivants :

a) the **back** of the room ;

b) put it on your **back** ;

c) the **back** of the chair.

La commission n'a pas jugé indispensable de retenir les deux dernières acceptions au niveau de la 6e.

La 2e acception apparaît en 5e, et la 3e est reportée à plus tard.
« Back » adverbe est retenu pour la liste de 6e.

2) On peut distinguer trois emplois usuels du verbe « ask » :

a) **ask** him a question ;

b) **ask** him to come ;

c) **ask** him for the key.

Seuls les deux emplois a) et b) ont été retenus pour la liste de 6e comme étant les plus urgents et les plus fréquents dans une situation d'apprentissage scolaire. L'emploi c) qui comporte une structure complexe à ce stade n'apparaît que dans la liste de 5e.

3) Il existe quatre acceptions pour le substantif « bath » :

a) have a **bath** (action de prendre un bain) ;

b) run a **bath** (le contenu) ;

c) clean the **bath** (le contenant) ;

d) a chemical **bath** (terme de chimie).

Seule l'acception a) a été retenue pour la 6e. Par contre les acceptions b) et c) n'ont été jugées indispensables ni en 6e, ni en 5e. Quant à d) son élimination est évidente à ce niveau.

Dans les limites d'acception et d'emploi ainsi précisées, la commission a aussi tenu compte du **pouvoir opératoire** des mots à retenir. Il est en effet important dans les débuts de l'apprentissage de la langue, au moment où l'on s'efforce de faire acquérir aux élèves les schémas fondamentaux et les mécanismes indispensables, de leur fournir les éléments lexicaux les plus aptes à la manipulation des structures les plus variées possibles. Ce pouvoir opératoire des unités lexicales thématiques est d'une qualité différente de celle des mots grammaticaux, sans le secours desquels il n'y aurait pas d'expression bien formée : c'est une qualité complémentaire dans la mesure où ces unités sont indispensables pour donner de la substance aux structures de base de la langue.

Exemple : utilité de « cheap » et d'« expensive » dès la 6e pour permettre de faire fonctionner le comparatif de supériorité.

2. Les mots grammaticaux (« G »)

Les mots grammaticaux correspondent au programme grammatical pour les classes de 6e et 5e à quelques exceptions près. Ils représentent une proportion importante de la liste de 6e : ce sont en particulier les mots-outils indispensables au bon fonctionnement fondamental des structures de la langue. Leur nombre est sensiblement moins élevé en 5e alors que le nombre de mots thématiques est nettement plus important qu'en 6e.

Parmi les adverbes, seuls figurent ceux qui posent un problème grammatical (ordre des mots, fonctionnement de la postposition). Les autres adverbes sont considérés comme des unités lexicales ordinaires. En ce qui concerne les verbes, n'apparaissent sous cette rubrique que les auxiliaires modaux retenus. Les verbes irréguliers proposés, dont la charge sémantique et la valeur thématique sont en général très élevées, ne figurent pas dans la liste des mots grammaticaux mais figurent dans la liste thématique.

NOTE :

Les catégories 1 et 2 (mots thématiques et mots grammaticaux) correspondent aux deux catégories d'unités lexicales dont il faut tenir compte pour apprécier la charge d'apprentissage du vocabulaire « actif » qui doit être parfaitement assimilé (c'est-à-dire mobilisable à chaque instant par l'élève à des fins **d'expression** et non pas seulement de compréhension).

3. Les expressions toutes faites ou « formulaic utterances » (« F ») (1)

Ce sont des expressions usuelles, « idiomatiques », de la conversation courante (exclamations, interjections, interpellations, excuses, salutations, remerciements).

Exemples : **all right, thanks, of course.**

4. Les séries (« S »)

Il s'agit des mots qui servent constamment à indiquer le jour, la date, la page d'un livre ou le numéro d'un exercice. Les quatre séries retenues sont :

- — les jours de la semaine,
- — les mois,
- — les nombres cardinaux : de 1 à 100 en 6° ; au-delà en 5°,
- — les nombres ordinaux : de 1er à 31e (pour l'expression de la date) en 6° et au-delà en 5°.

5. Pédagogie (« P »)

Il s'agit des mots et expressions qui servent constamment à donner des instructions ou des encouragements, à désigner quelques objets usuels de la classe. Tout comme les mots appartenant à des séries, leur apprentissage se fait grâce à leur fréquente manipulation dans les échanges entre professeurs et élèves. Ces mots et expressions sont imprimés en italique.

B - REGLES D'INCLUSION DES UNITES LEXICALES RETENUES

1) *Figure dans ces listes toute unité lexicale (lexème simple ou complexe) qui se différencie d'une autre par*

a) sa catégorie grammaticale,

ex. : **after,** préposition, différent de **after**, conjonction,

---

(1) Expression empruntée à la terminologie de Randolph Quirk (A Grammar of Contemporary English), voir bibliographie, p. 46.

b) sa morphologie : **men** différent de **man** ; **break** de **broke** et **broken**,

c) son contenu sémantique,

ex. : **right** (≠ left) différent de **right** (≠ **wrong**).

Les « mots » ainsi définis portent une « charge d'apprentisage » plus ou moins importante. La commission s'est essentiellement donné pour tâche de répartir cette charge de façon rationnelle.

2) *Deux exceptions à la règle* 1

a) Les substantifs dont le pluriel graphique se forme par simple adjonction de « s » ou « es » n'ont qu'une entrée.

b) Les flexions « s » (ou « es ») « ed » et « ing » ne donnent pas lieu à des entrées séparées.

3) *Cas des auxiliaires et des modaux*

a) Ils obéissent à la règle 1) :

Ex. : entrées distinctes pour **am, are, is**, etc.,
had et **have,**
may et **might**, etc.

b) Lorsque la négation « not » est intégrée à une forme verbale sous sa forme réduite (« n't »), le mot figure séparément dans la liste :

Ex. : **aren't ; can't ; isn't ; don't ; won't** ; etc.

c) Lorsqu'un auxiliaire a une forme pleine et une forme réduite, elles sont données ensemble sous la même entrée :

Ex. : **am, 'm ; is, 's ; shall, 'll ; will, 'll** ; etc.

4) *Cas des verbes composés*

On a distingué par la typographie trois cas, suivant la façon dont la particule est liée au verbe pour en modifier plus ou moins le sens :

a) Si le verbe composé que l'on peut créer par juxtaposition du verbe et d'un adverbe postposition figurant dans la liste (**back, in, out**) a un sens premier qui résulte de la simple addition des sens des deux mots juxtaposés, ce verbe n'apparaît pas dans la liste comme unité lexicale indépendante.

b) Si le sens du verbe intransitif ne résulte pas de la simple addition de la postposition, ce verbe est traité comme une unité lexicale indépendante.

Ex. : **get up** (v. intr.) T ; en effet le sens de « get up » ne résulte pas de la simple addition de « get » (à la charge sémantique faible, du fait de sa polysémie) et de la postposition « up ».

c) Les verbes transitifs dont les deux éléments peuvent être parfois séparés comportent le signe + entre les deux éléments :

Ex. : **get** + **out** (v. tr.) (« get them out »).

5) *Cas des noms composés*

La commission n'a retenu que ceux dont l'apprentissage lui paraissait indispensable au niveau considéré (ex. : **bathroom** en 6e, **dining-room** en 5e) étant entendu qu'on aura toute latitude pour enseigner dans des limites raisonnables d'autres noms composés qui résultent de la simple addition des contenus sémantiques de deux unités lexicales qui figurent dans la liste (ex. : **apple-tree**).

C - PRESENTATION DES LISTES

I - Ordre des mots

Les deux listes sont présentées sur deux colonnes dans un ordre alphabétique unique. Une troisième colonne donne les exemples d'emploi de certains mots, suivant les principes indiqués plus bas. Les unités lexicales complexes (locutions ou expressions) apparaissent dans l'ordre alphabétique du mot principal :

Ex. : **able (be...).**

(Ce n'est pas l'adjectif **able** qui doit être enseigné, mais l'expression **be able** sous ses différentes formes, qu'il est plus commode de trouver à la lettre A).
Pour les expressions toutes faites cependant (F), c'est l'ordre alphabétique du permier mot de l'expression qui a été retenu :

Ex. : **all right.**

II - Catégorie des mots

Les cinq catégories sont indiquées par des lettres majuscules.

T : mots **thématiques** ou à sémantisme plein ;

G : mots **grammaticaux** (1) ;

S : **séries** ;

F : « **formulaic** utterances » (ex. **all right...**) ;

P : **pédagogie.**

(1) Voir p. 32 la liste des abréviations utilisées.

III - Exemples d'emploi

Les exemples (un ou plusieurs suivant le cas) ne sont fournis en principe que pour certains mots **thématiques** ou **grammaticaux** (T ou G). Ce ne sont pas des phrases types.

Ils ont pour but de limiter les acceptions ou les constructions à enseigner obligatoirement.

Ex. agree (v. intr.) **I quite agree with you.**

Les autres acceptions et constructions n'entrent pas dans le minimum lexical et sont donc laissées à la discrétion des professeurs et des auteurs de manuels.

Aucun exemple n'est donné dans les cas suivants :

1) **Mots appartenant aux catégories « S », « F », « P » :** leur champ sémantique est en principe limité et précis ou indiqué par la situation pédagogique.

Il n'est fait exception à cette règle que pour introduire une distinction jugée nécessaire.

Ex. : get + out (v. tr.) P - **Get out your books.**

2) **Mots grammaticaux** (G) quand la précision que l'exemple aurait apportée serait liée à un problème de grammaire plus que de lexique.

Ex. : **a, an** (art.) G.

Par contre : about (prep.) G **What's the lesson about ?**

3) **Mots thématiques** (T) ayant **une seule acception.**

Ex. : **airport** (n.) T.

Ou ayant **plusieurs acceptions également utiles** au niveau considéré.

Ex. **hall** (n.) T.

Par contre address (n.) T **Write the address on the letter.**

IV - Abréviations, signes et conventions

Dans la liste de 5e, l'astérisque désigne une unité lexicale qui se trouve déjà dans la liste de 6e sous la même forme graphique.

Cette unité n'entraîne qu'une charge d'apprentissage réduite. Il peut s'agir :

1) du participe passé d'un verbe dont le prétérit, de forme identique a déjà été donné :

ex. : **caught** (v.) pret T * **caught** (v.) p.p. T (1)

---

(1) Etant donné que le *Programme grammatical* ne prévoit l'étude du « present perfect » qu'en 5e, de même que l'étude méthodique du passif, seuls ont été retenus en 6e les participes passés des verbes irréguliers du programme de 6e ayant un emploi adjectival : ex. *shut*.

Les autres n'apparaissent que dans la liste de 5e.

2) de la réapparition du même mot dans une autre catégorie lexicale ou grammaticale :

ex. :	**answer** (v.)	P	* **answer** (v. tr. et intr.)	T	
			* **answer** (n.)	T	
ex. :	**after** (prép.)	G	* **after** (conj.)	G	

3) de la réapparition du même mot dans une autre construction :

ex. :	**as** (adv.)	G	* **as** (conj. temps)	G	
	as (conj.)	G			
	ask (v. tr.)	T	* **ask** (v. tr. et intr.)	T	**Ask your mother for the key**
	Ask him a question				
	Ask him to come				

4) de la réapparition du même mot dans une autre acception :

ex. :	**country** (n.)	T	* **country** (n.)	T	**How many countries have you visited ?**
	He lives in the country				
	do (aux.)	G			
	do (v. tr.)	T	* **do** (v. tr.)	T	**Do your room now**
	What are you doing ?				

Les abréviations entre parenthèses indiquent la catégorie grammaticale des mots des catégories T et G. Elles ne sont données pour les mots de la catégorie Pédagogie que pour éviter une ambiguïté.

Ex. : **answer** (v.) P.

On trouvera à la page suivante la liste alphabétique des abréviations utilisées.

Liste alphabétique des abréviations

adj.	adjectif
adv.	adverbe
art.	article
aux.	auxiliaire
aux. mod.	auxiliaire modal
conj.	conjonction
excl.	exclamatif
inter.	interrogatif
loc. prép.	locution prépositionnelle
loc. v.	locution verbale
n.	nom
nég.	négatif
post.	postposition
p.p.	participe passé
prép.	préposition
prét.	prétérit
pron.	pronom
rel.	relatif
v. (1)	verbe
v. intr.	verbe intransitif
v. tr.	verbe transitif

Voir la liste publiée au B.O.E.N. du 28-9-72

Circulaire n° 72 345 du 25 septembre 1972

(1) Soit :
— copule,
— auxiliaire,
— forme irrégulière d'un verbe,
— en pédagogie.

Annexe I

LISTE ANALYTIQUE DES OUVRAGES SCIENTIFIQUES DÉPOUILLÉS

Chacun des ouvrages ci-dessous est indiqué par un symbole letire sur les fiches (1).

1 - A GENERAL SERVICE LIST OF ENGLISH WORDS (W)

With Semantic Frequencies and a Supplementary Word-list for the Writing of Popular Science and Technology compiled and edited by Michael West, M.A., D. Phil.

Longmans, 1936, 588 pages, Revised and enlarged edition ; first published 1953, Ninth impression 1965.

(corpus de langue écrite de 4 millions 1/2 de mots, critère de fréquences groupées : mots de base et dérivés, liste de 2 000 mots de base).

Cet ouvrage représente la révision du **Report on Vocabulary Selection,** établi par un groupe de spécialistes en 1936, à partir de l'étude de I. Lorge et E. L. Thorndike, **A Semantic Count of English Words** (1938). Les 2 000 mots considérés comme la base du vocabulaire de l'anglais langue étrangère sont accompagnés d'un pourcentage d'emploi des différentes acceptions.

---

(1) Voir annexe II.

Ainsi **GAME** est affecté du nombre **638** parce qu'il a été trouvé 638 fois (1). Le sens **amusement, children's play** est apparu dans 9 % des cas ; dans 38 % des cas c'est « **with the idea of competition as in football, card games** » etc. Un signe (?) suggère que le sens **animals** (11 %) peut ne pas être retenu par le professeur. Un autre signe ([ ]) indique que le sens **fun** (make game of) ne devrait pas être enseigné.

2 - THE TEACHER'S WORD BOOK OF 30,000 WORDS (T 1 - T 2 - T 3)

By Edward L. Thorndike and Irving Lorge

Bureau of Publications
Teachers College, Columbia University, New York, 1944, 274 pages,
3e *édition,* 1959

(Corpus de langue écrite, 18 millions de mots, critère de fréquence).

Ce livre est une édition augmentée du **Teacher's Word Book** de Thorndike (20 000 mots) publié en 1931, qui était lui-même une extension de son premier **Teacher's Word Book** de 1921.

Pour chacun des 20 000 premiers mots retenus, les auteurs indiquent le coefficient de fréquence dans quatre corpus de 4 millions 1/2 de mots chacun, désignés par une lettre :

T - corpus du **Teacher's Word Book** de Thorndike de 1931 (livres scolaires, Bible et un certain nombre de classiques) ;

L - corpus de magazines et journaux dépouillés par Lorge ;

J - corpus de 120 livres pour la jeunesse dépouillés par Thorndike (« juvenile books ») ;

S - corpus utilisé par Lorge et Thorndike pour leur « semantic count ».

Ils en déduisent le coefficient de **fréquence générale** du mot (G) en divisant le total des quatre coefficients par 18, ce qui donne le nombre d'occurences par million de mots. Dans cette liste alphabétique de 19 440 mots, tout coefficient compris entre 50 et 99 est représenté par la lettre A et tout coefficient égal ou supérieur à 100 par la double lettre AA (soit 1 069 mots affectés de la double lettre AA et 952 de la lettre A).

---

(1) Le coefficient de fréquence le plus élevé est celui de la conjonction « and » ; il est de 106 064.

Parmi les mots thématiques, on peut citer aux deux extrémités :

see (v.)	7 048	heap (n.)	76
take (v.)	7 008	fork (n.)	70
part (n.)	6 945	damp (adj.)	67.

	G	T	L	J	S
Ex. : GAME	AA	700	966	700	639

Les auteurs donnent ensuite, en deux listes supplémentaires, les quelque 10 000 mots ajoutés à la première liste de 1931 :

1) une liste de 9 202 mots apparaissant moins d'une fois sur 1 million mais plus de 4 fois sur 18 millions ;

2) une liste de 1 358 mots apparaissant 4 fois sur 18 millions.

Ils donnent enfin, c'est sans doute le résultat le plus intéressant pour une exploitation pédagogique :

1) la liste alphabétique des mots les plus fréquents dans les corpus L et S (coefficient égal ou supérieur à 1 000) avec l'indication, toutes les fois qu'ils sont connus, de leurs deux coefficients.

(le mot « game » n'apparaît pas dans cette liste).

2) deux listes alphabétiques correspondant aux deux premières tranches de 500 mots parmi les 1 000 mots les plus fréquents (ceux qui sont affectés des lettres AA dans la liste générale) (1).

ex. : GAME apparaît dans la 2e tranche.

3 - THE WORDS AND SOUNDS OF TELEPHONE CONVERSATIONS (F)

By Norman R. French, Charles W. Carter, Jr. and Walker Koenig.

Bell System Technical Journal 1930. 34 pages, 79 390 mots.

(Corpus de langue parlée, critères de fréquence et de répartition).

Cet article contient une étude sur le vocabulaire (et la fréquence relative des sons de la langue parlée) relevé dans 500 conversations téléphoniques choisies elles-mêmes parmi 1 900 conversations entre 3 800 différents locuteurs américains. (Dans 90 % des cas il s'agissait de conversations d'affaires).

La liste de mots la plus significative groupe ceux qui apparaissent dans au moins 5 conversations sur les 500 analysées, soit 737 mots sur un total de 2 240 mots différents.

---

(1) Sur la fiche établie par la commission, T1 représente la 1re tranche de 500 mots, T2 la 2e tranche et T3 les 1 000 mots suivants (affectés de la lettre A dans la liste générale).

Dans la colonne A de la liste par ordre de fréquence décroissante, on trouve le nombre total d'occurences de ce mot (fréquence), dans la colonne B le nombre de conversations dans lesquelles ce mot a été utilisé (répartition).

A fréquence		B répartition
3 990 (maximum)	I	467
3 540	you	499 (maximum)
3 110	the	496
....	...	...
		etc.
9	game	8
5 (minimum)	crowd	5 (minimum)

4 - TEACHING ENGLISH IN DIFFICULT CIRCUMSTANCES (W)

Teaching English as a Foreign Language, with notes on the techniques of textbook construction, by Michael West, M.A. D.Phil. Longman. First published 1960, 4 *th impression,* 1966. 136 *pages.* (*Corpus de langue parlée, critères fonctionnels multiples, liste de* 1 200 *mots.*)

Dans ce livre qui est né de quarante années de réflexion sur l'enseignement de l'anglais langue étrangère, il a été retenu surtout une présentation d'un vocabulaire minimum de la langue orale : **The Minimum Adequate Speech Vocabulary,** qui comprend 1 200 mots, ainsi que sa classification en sept groupes :

A) The Earth, B) The Self, C) The Home, D) The Intellect, E) Business, F) Relaxation et G) Public Life.

Chacun de ces groupes contient plusieurs sous-titres, en tout 40.

Ex. : F) Relaxation

31) Games, Athletics and Sport.

To aim, amuse, arrange, ball, basket, basket-ball, bicycle, boating, captain, cards, climb, dance, fair/unfair, fast, field, football, game etc. (au total 73 mots sous ce paragraphe 31).

5 - A STUDY OF THE ORAL VOCABULARY OF ADULTS (S)

An Investigation into the Spoken Vocabulary of the Australian Worker by F. J. Schonnell, I. G. Meddleton, B. A. Shaw (1956).

University of London Press Ltd, Warwick Square London EC4, 176 pages. (Corpus de langue parlée par des ouvriers australiens, 1/2 million de mots, critère de fréquences groupées, liste de 1 000 mots de base).

Il s'agit du résultat d'un travail d'équipe entrepris en 1951 en vue de l'enseignement d'un vocabulaire minimum aux immigrants venus d'Europe et du Commonwealth. Un 1/2 million de mots, utilisés par 3 000 ouvriers, hommes et femmes, dans quelque 1 500 situations différentes, ont été recencés et classés.

La liste comprend les 1 000 premiers mots-chefs de file (« head-words ») dans l'ordre alphabétique, avec leur indice de fréquence et suivis de tous les mots de la même famille qui ont été recencés avec leur propre indice. En totalisant ces indices, on obtient l'indice de fréquence de la famille.

ex : GAME	(n.)	126
	games (n.)	25
	game (n.)	1
	game's (is)	1
		153 INDICE de fréquence de la famille de game.

L'inconvénient évident de cette liste, est que les différents sens d'un même mot ne sont pas toujours indiqués.

Ex. : Game.

6 - L'ELABORATION DU FRANÇAIS FONDAMENTAL (1er DEGRE) (FF 1 - FF 2)

Etude sur l'établissement d'un vocabulaire et d'une grammaire de base par G. Gougenheim, P. Rivenc, R. Michéa, A. Sauvageot. Didier, Paris, 1964.

[Corpus de langue parlée, critères de fréquence, répartition et disponibilité, liste de 1 475 mots (1er degré) complétée ultérieurement par une deuxième liste de 1 609 mots (2e degré).]

Les recherches qui ont abouti à l'établissement du français fondamental (1er degré) sont intéressantes à un double titre : pour la première fois on est parti d'un corpus de langue parlée ; les limitations du critère de fréquence ont d'autre

part permis de découvrir un critère complémentaire ; le critère de disponibilité (1).

Pour évaluer **le degré de disponibilité** des mots, les auteurs de cette étude ont distingué seize centres d'intérêt :

1) les parties du corps ;
2) les vêtements ;
3) la maison, etc.

A l'intérieur de chacun de ces centres d'intérêt, les mots jugés les plus utiles ont été retenus. Cette méthode a inspiré une étude comparable sur le vocabulaire anglais (voir n° 8 ci-dessous).

La liste de fréquence établie par la commission du Français Fondamental comportait au départ 1 063 mots allant de la fréquence 14 083 à la fréquence 20 (le coefficicent de répartition allant de 163 à 1).

En décidant de ne pas descendre au-dessous de la fréquence 29 et pour éliminer un certain nombre de mots qu'il ne paraissait pas opportun d'inclure dans la liste définitive, on n'a finalement retenu que 701 mots parmi les plus fréquents. Sur les 1 475 mots du Français Fondamental 1er degré plus de la moitié ont été fournis par les enquêtes sur les mots disponibles et par une décision de la commission de donner à sa liste un caractère éducatif, universel et utile.

Les mots du Français Fondamental (1er degré) ont été traduits en anglais par notre commission et reclassés par ordre alphabétique. La commission a en effet considéré qu'il serait intéressant de tenir compte des résultats obtenus par ces recherches sur le vocabulaire français. Le coefficient de fréquence a été porté sur les fiches toutes les fois qu'il était donné.

7 - CRITERES DE CHOIX DES ELEMENTS D'UN « CORPUS » POUR LA SIXIEME (Ba)

par BAILLY Danièle

(*Problèmes de contenu d'un manuel*), 1967.
(*Corpus de langue parlée établi à partir des structures fondamentales, critères multiples, liste de 600 mots.*)

L'auteur a tenté d'analyser les critères de choix des éléments grammaticaux et lexicaux qui devraient figurer dans un manuel d'enseignement au niveau de la classe de sixième.

Parmi ces critères, D. Bailly a considéré le critère grammatical comme fondamental. En effet, une des finalités majeures de l'enseignement de l'anglais semble être à ses yeux la mise en évidence, à l'intention des élèves, des struc-

(1) Voir pages 9 et 10.

tures de la langue, dont l'appropriation ouvre la voie à une réflexion sur le langage en général.

Viennent ensuite : le critère psycho-pédagogique (processus de l'apprentissage, fixation des structures, etc.), le critère culturel, le critère d'efficacité dans le maniement de la langue (possibilités pour l'élève de s'exprimer de manière autonome dans des situations différentes des situations proposées dans la méthode, le critère psychologique proprement dit (motivations de l'enfant, appréhension du réel, etc.).

Une liste de mots s'est donc ainsi constituée, dont chacun répond à des exigences de « rentabilité » sur le plan à la fois linguistique, psychologique et proprement pédagogique.

Exemple : pour le mot « game »

- Intérêt grammatical : mot « dénombrable » d'emploi régulier.
- Intérêt psycho-pédagogique : champ sémentique différent en anglais et en français.

(Pas d'interférence grammaticale, soit avec le français, soit à l'intérieur même de l'anglais. Emploi contingent dans les exercices structuraux).

- Intérêt d'efficacité : référence éventuelle à la situation (la classe, la vie).
- Intérêt psychologique : assez grand (cf. motivation, etc.).
- Appréciation de synthèse : terme d'une exploitation facile, d'une rentabilité moyenne, n'illustrant pas spécifiquement une structure fondamentale.

Ce mot est à retenir surtout pour son insertion possible dans un thème motivant pour l'élève.

8 - BUNTTÚS GAEILGE RÉAMTHUJARASCÁIL (Ó)

Colmán L. Ó'HUALLACHÁIN.

Etude de disponibilité : vocabulaire anglais d'enfants irlandais, 1966, 207 *pages.*

(*Corpus de langue parlée par des enfants irlandais, critère de disponibilité, liste de* 1 100 *mots.*)

L'intérêt majeur de la liste donnée en fin de cet ouvrage (rédigé en langue gaélique) est la classification en vingt centres d'intérêt des quelques 1 100 mots retenus (noms, verbes, adjectifs).

Chacune des 20 listes ne comporte que les mots ayant un coefficient de disponibilité égal ou supérieur à 20, le maximum étant 330.

Exemple : Centre d'intérêt n° 2 - **Sports** : (50 mots) football (188), running (139), hurling (126), tennis (122), jumping (106), run (92), boxing (84), golf (80), * swimming (79), hockey (78), jump (78), rugby (75), soccer (64), cycling (62), basketball (60), cricket (52), play (51), racing (50), camogie (48), ball (47), drill (47), kick (44), walking (43), wrestling (43), swim (42), handball (40), skipping (39), healthy (35), win (34), * swimming (33), team (30), hurley (29), game (28), fishing (27), netball (27), horseracing (25), score (25), training (25), walk (25), high jump (24), skating (24), dancing (23), train (23), weight-lifting (23), throw (21), hurl (20), race (20), riding (20), skiing (20), skip (20).

A titre indicatif voici les 6 mots qui ont le plus grand coefficient dans l'ensemble de la liste :

minister 330 (centre d'intérêt : politique) ;

summer 229 (centre d'intérêt : temps, saisons) ;

bus 221 (centre d'intérêt : transports) ;

knife 220 (centre d'intérêt : repas) ;

desk 216 (centre d'intérêt : école) ;

bread 202 (centre d'intérêt : repas) ;

(*) Le mot « swimming » est donné deux fois dans la liste.

Annexe II

MODÈLE DE FICHE

GAME n	w 638

6°	5°	
X	X	CF
X		GM
	X	LK
	X	LLD
X		PP
X		RH

Sens

1) with the idea of competition (a ... of football) 38 %

2) a particular contest (playing a poor...) 23 %

3) amusement, children's play 9 %

X		CGG
	X	FW

4)

	X	G
	X	H

5)

(1)	(2) X	(3)	T
X	F		
X	W		
X	S	152	
X	FF 1	jeu n	
	FF 2		
X	Ba		
X	0	56	

CONTENU DE LA FICHE

1. L'unité lexicale est portée en haut à gauche avec l'indication de sa catégorie grammaticale.

2. Le numéro dans l'angle droit de la fiche : W ..... indique le coefficient de fréquence dans WEST : « A General Service List of English Words ».

3. Les différents sens du mot sont indiqués selon le West (General Service List) et dans leur fréquence décroissante, en pourcentage.

4. Les sigles correspondent aux manuels et ouvrages suivants :

CF	Carpentier-Fialip - L'anglais vivant (éd. rouge), 6°-5° - Hachette.
GM	Guitard-Marandet - Cours d'anglais - Nathan, 6°-5°.
LK	Laffay-Kerst - England - Masson, 6°-5°.
LLD	Lebettre-Lonjaret-Denis - Parlons anglais - Bourrelier, 6°-5°.
PP	Perthuisot-Papy - English ways - Hatier, 6°-5°.
RH	Richard-Hall - L'anglais par l'illustration - Hachette, 6°-5°.
CGG	Capelle-Garnier-Girard - Junior Course. Passport to English - I et II - Didier, éd.
FW	Filipovic-Webster - Saint-Cloud Zagreb - British English by the audio-visual method - I et II - Didier.
G	Gatenby - A Direct Method English Course - I et II - Longman.
H	Hornby - Oxford Progressive English Course - I et II - O.U.P. ed.
T	Thorndike and Lorge, The Teacher's Word Book of 30 000 Words - Columbia University, éd.
F	French-Carter-Koenig - The Words and Sounds of Telephone Conversations - Bell System Technical Journal 1930, éd.
W	West - Teaching English in Difficult Circumstances - Longman, éd.
S	Schonell - A study of the Oral Vocabulary of Adults-University of London Press, éd. (1).
FF 1	Français Fondamental, premier degré.
FF 2	Français Fondamental, deuxième degré.
Ba	Danielle Bailly - Fichier ronéotypé.
O	O'Huallachain (1).

(1) Le nombre porté à droite indique le coefficient de fréquence du mot.

Annexe III

LES MOTS GRAMMATICAUX

Le verbe		Le substantif et le groupe substantival				
Auxiliaires et modaux et locutions verbales		Les déterminants				Pronoms personnels et réfléchis
		Articles définis - indéfinis	Démonstratifs	Quantificatifs	Possessifs	
6e	5e	6e	6e	6e	6e	6e
am, 'm are aren't be can cannot, can't could couldn't did didn't do does doesn't don't — has hasn't have, 've is, 's isn't must mustn't shall, 'll shan't was wasn't were weren't will, 'll won't there is (1)	able (be) allowed (be) been going (be) had have got let may might must needn't should, 'd shouldn't would, 'd wouldn't	a an the ∅	that these this those	all a lot of another any every lots of many more (not) much no one other others plenty of some	's her his its my our their your	he her him I it me she them they us we you
				5e	5e	5e
				anybody anything both everybody everything many (as... as) many (not { so... as / as... as }) many too most much (as... as) much (not { as... as / so... as }) much (too) nobody nothing somebody something	hers his mine ours theirs yours	herself himself itself myself ourselves themselves yourself yourselves

(1) « is » et autres formes de ce prédicat d'existence retenues pour la classe de sixième.

GRAMMATICAUX

La phrase

Relatifs	Coordination	Subordination	Adverbes (lieu, temps, interrogatifs, quantité, etc.)		Prépositions	Adjectifs et pronoms interrogatifs
6e	6e	6e	6e	5e	6e	6e
which who	and but or	as (... as) because if than that when where	again all always as away (postposition) back (postposition) before better down (postposition) far how how many (int.) how much (int.) in (postposition) more much near never not, n't now off (postposition) often on (postposition) only over out (postposition) so sometimes then too up (postposition) very when where why worse	about already also anywhere bit (a) either how far how long how many how much (excl) indeed inside instead instead of just much (as... as) much (not { as... as / so... as }) much (too) nearly outside over round somewhere soon to (but) well what (excl) yet	about after at before behind between by down for from in in front of into near of on out of over past to under with	what who whose
5e		5e			5e	5e
whom whose		after as (tps)			across all over along like round though without	which

BIBLIOGRAPHIE

1. AUSTIN (J. L.). — **How to do things with words** (1962).

2. BAILLY (Danielle). — **L'enseignement de l'anglais à des élèves débutants, classe de 6e - Appropriation grammaticale et opérations psycholinguistiques** - Thèse de Doctorat de 3e cycle - Université de Paris VII (U.E.R. d'Anglais - Charles V), mars 1971.

3. COFFER (Charles N.). — **On some factors in the organizational characteristics of free recall** - *American Psychologist* 20 (1965), p. 261-272, repris dans *Readings in the Psychology of Language* publié par Leon A. Jacobovits and Murray S. Miron-Prentice Hall Psychology series, p. 587-603.

4. DAMERAU (Fred. J.). — **Statistical Word Association and Synonymy** - IBM Research Division, 1965.

5. FRENCH (N. R.), CARTER (Ch. W.), KOENIG (Jr. W.). — **The Words and Sounds of Telephone Conversations.** Bell System Technical Journal, 1930.

6. FRIES (C. C.) and TRAVER (A. A.). — **English Word Lists. A study of their adaptability for instruction** (Washington, D. C. ; Amer. Council of Educ., 1940).

7. GALISSON (R.). — **Inventaire thématique et syntagmatique du français fondamental** (Paris, Hachette-Larousse, 1971).

8. GALISSON (R.). — **L'apprentissage systématique du vocabulaire** (Paris, Hachette-Larousse, 1970).

9. GOUGENHEIM (G.), MICHEA (R.), RIVENC (P.), SAUVAGEOT (R.). — **L'élaboration du français fondamental (1er degré)** (Paris, Didier, 1964).

10. HIGA (M.). — **The psycholinguistic concept of « difficulty » and the teaching of foreign language vocabulary** (dans *Language Learning*, vol. XV, nos 3 et 4, 1965).

11. DAVIS HOWES. — **A Word Count of Spoken English,** in *Journal of Verbal Learning and Verbal Behaviour* 5 (pp. 572-604), 1966.
(250,000 words of recorded interviews with university students and hospital patients. Data for sub-samples of 100,000 words each from student population. 9,699 words in all.).

12. KUCERA and FRANCIS. — **Computational Analysis of Present Day American English** (Brown University Press).
Student population. 9,699 words in all.
(500 samples of 2,000 words from 14 different types of printed source.)

13. LADO (R.). — **Linguistics across Cultures** (Ann Arbor, Michigan University Press, 1957).

14. MACKEY (W. F.). — **Language Teaching Analysis** (Londres, Longman, 1965).

15. MACKEY (W. F.) and SAVARD (J. G.). — **The indices of coverage** (dans *IRAL*, vol. V, 2-3, 1967).

16. MELTON (A. W.). — **Implications of short-term memory for a general theory of memory,** in *Journal of Experimental Psychology*, vol. 62 (1961), p. 335-347.

17. O'HUALLACHAIN (C. L.) — **Buntûs Gacilge** - Reamhthuarascáil 1966 -

18. **Programme grammatical par l'enseignement de l'Anglais.** — SEVPEN, 1965.

19. QUIRK (R.), GREENBAUM (S.), LEECH (G.) and SVARTVIK (J.). — **A Grammar of Contemporary English** (Londres, Longman, 1972).

20. SAVARD (J. G.). — **La valence lexicale** (Paris, Didier, 1970).

21. SAVARD (J.-C.) et RICHARDS (Jack). Préface de Mackey. — **Les indices d'utilité du vocabulaire fondamental français,** Les presses universitaires de l'Université de Laval.

22. SCHONELL (F. J.), MEDDLETON (I. G.), SHAW (B. A.) — **A Study of the Oral vocabulary of Adults.** Research study 1. University of Queensland Press and University of London Press Ldt, 1965.

23. SINCLAIR (J. M.). — **Beginning the Study of Lexis in « In Memory of J. R. Firth »** (Longmans, 1966).

24. SINCLAIR (J. M.) (with S. JONES and R. DALEY). — **English Lexical Studies** (OSTI Report Jan. 1970), Dept of English, The University, Birmingham, England.

25. THORNDIKE (E. L.) et LORGE (I.). — **The Teacher's Word Book of 30,000 words** (Columbia University, 1959).

26. WEST (M.). — **A General Service List of English Words**, Longman, 1965.

27. WEST (M.). — **Teaching English in Difficult Circumstances**, Longman, 1960-1966.

28. ZIPF (G. K.). — **The Psycho-Biology of Language**, 1931.

29. THE AMERICAN HERITAGE WORD FREQUENCY BOOK
(Published by :

The American Heritage Publishing Co.
551 Fifth Avenue, New York
NY 10017 USA.)

(5 millions words of running text, 500-word samples drawn from « the range of required and recommended reading to which students are exposed in grades 3 through 9 in the US.)

Dépôt légal n° 4 287 - 2e trim. 1973 - Imp. BIALEC - NANCY

Les publications de l'I.N.R.D.P.

(dont le catalogue gratuit vous sera envoyé sur demande)

sont diffusées par le S.E.V.P.E.N. à

AJACCIO - C.D.D.P. : Avenue Pugliesi - Conti
ALBI - C.D.D.P. : Centre administratif - rue du Général Giraud
ALENÇON - C.D.D.P. : Place Sonet
AMIENS - C.R.D.P. : 33, rue des Minimes
ANGERS - C.D.D.P. : 9, rue Grandet
ANGOULÊME - C.D.D.P. : 1, rue Vauban
ARRAS - C.D.D.P. : 39, rue aux Ours
AUCH - C.D.D.P. : rue Boissy-d'Anglas
AVIGNON - C.D.D.P. : 8, rue Frédéric-Mistral
BEAUVAIS - C.D.D.P. : 22, avenue Victor-Hugo - B.P. 321
BESANÇON - C.R.D.P. : 16 et 17, rue Ernest-Renan - B.P. 78
BORDEAUX - C.R.D.P. : 75, cours d'Alsace-Lorraine
BOURG-EN-BRESSE - C.D.D.P. : 6, rue Jules-Ferry
BOURGES - Antenne du C.R.D.P. : 9, rue Edouard-Branly
BREST - C.D.D.P. : 108, rue Jean Jaurès
CAEN - C.R.D.P. : 21, rue du Moulin au Roy
CAHORS - Antenne du C.R.D.P. : Cité Administrative - Quai Cavaignac
CARCASSONNE - C.D.D.P. : 56, avenue du Docteur Henri-Goût
CHALONS S/MARNE - Antenne du C.R.D.P. : Cité administrative - Caserne Tirlet
CHAUMONT - C.D.D.P. : rue Haeusler
CHARLEVILLE / MÉZIÈRES - C.D.D.P. : 18, rue Voltaire
CLERMONT-FERRAND - C.R.D.P. : 15, rue d'Amboise
DIGNE - C.D.D.P. : Ancien Lycée de jeunes filles - Place des Cordeliers
DIJON - C.R.D.P. : Campus Universitaire de Montmuzard, Bld Gabriel - B.P. 490
FOIX - C.D.D.P. : 31 bis, avenue du Général de Gaulle
FORT-DE-FRANCE (Martinique) - C.D.D.P. : 49, rue Victor-Sévère
GAP - Antenne du C.R.D.P. : 4, avenue du Maréchal Foch
GRENOBLE - C.R.D.P. : 11, rue du Général Champon
LAON - C.D.D.P. : rue Ferdinand-Thuillard - Impasse de l'Église
LA ROCHELLE - C.D.D.P. : 19, rue Vaugoin
LE MANS - C.D.D.P. : 31, rue des Maillets
LILLE - C.R.D.P. : 3, rue Jean-Bart
LIMOGES - C.R.D.P. : 44, rue Gay-Lussac
LYON - C.R.D.P. : 47, rue Philippe-de-Lassalle (4e)
MARSEILLE - C.R.D.P. : 55-57, rue Sylvabelle (6e)
MONTAUBAN - C.D.D.P. : 5, rue du Fort
MONT-DE-MARSAN - C.D.D.P. : Cité Galliane - Avenue Cronstadt
MONTPELLIER - C.R.D.P. : allée de la Citadelle
NANCY - C.R.D.P. : 99, rue de Metz
NANTES - C.R.D.P. : 17, rue Gambetta - B.P. 1001
NEVERS - C.D.D.P. : École du Château
NICE - C.R.D.P. : 117, rue de France
NIMES - C.D.D.P. : 10, Grand'Rue
ORLÉANS - C.R.D.P. : 55, rue Notre-Dame-de-Recouvrance
PARIS - C.R.D.P. : 29, rue d'Ulm (5e) - Librairie : 13, rue du Four (6e)
PAU - C.D.D.P. : villa Nitot - avenue Nitot - B.P. 299
PERPIGNAN - C.D.D.P. : 24, rue Émile-Zola
POINTE-A-PITRE (Guadeloupe) - C.D.D.P. : rue du Gouverneur Félix-Éboué
POITIERS - C.R.D.P. : 6, rue Sainte-Catherine
REIMS - C.R.D.P. : 47, rue Simon
RENNES - C.R.D.P. : 92, rue d'Antrain
RODEZ - C.D.D.P. : ENI, 12, rue Sarrus
ROUEN - ST-CLÉMENT - C.R.D.P. : Cédex 3038 - 2, rue du Docteur Fleury Le Mont-Saint-Aignan
SAINT-DENIS (la Réunion) - C.D.D.P. : 12, rue de la Victoire
SAINT-ÉTIENNE - C.D.D.P. : 16, rue Marcellin-Allard
STRASBOURG - C.R.D.P. : 5, Quai Zorn - B.P. 279/R7
TARBES - C.D.D.P. : rue Georges-Magnoac - B.P. 205
TOULOUSE - C.R.D.P. : 3, rue Roquelaine
TOURS - C.D.D.P. : 1, rue Gutenberg
TROYES - C.D.D.P. : 1, rue Bégand
VALENCE - C.D.D.P. : "Ancienne Préfecture" - Place Le Cardonnet - B.P. 498

INSTITUT NATIONAL
DE RECHERCHE
ET DE DOCUMENTATION
PÉDAGOGIQUES

inrdp

SERVICE D'ÉDITION ET DE VENTE
DES PUBLICATIONS
DE L'ÉDUCATION NATIONALE
29 RUE D'ULM
75 230 PARIS - CEDEX 05